UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/008,574	10/26/2001	D. Wade Walke	LEX-0264-USA	6643

24231 7590 07/13/2004

LEXICON GENETICS INCORPORATED
8800 TECHNOLOGY FOREST PLACE
THE WOODLANDS, TX 77381-1160

EXAMINER	
MURPHY, JOSEPH F	

ART UNIT	PAPER NUMBER
1646	

DATE MAILED: 07/13/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/008,574	WALKE ET AL.	
	Examiner	Art Unit	
	Joseph F Murphy	1646	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *04/26/2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1 and 3-18* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1, 3-18* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *3/24/03 3/29/02*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☒ Other: *Sequence Comparison A*.

DETAILED ACTION

Election/Restrictions

Applicant's election with traverse of Group II, claims 4-5 in the response filed 04/26/2004 is acknowledged. The traversal is on the ground(s) that the nucleotide sequence of SEQ ID NO: 3 is a fragment of the nucleotide sequence of SEQ ID NO: 1, and that the claims should be searched together. In response to this argument, the groups will be examined together. Claims 1, 3-18 are pending and under consideration.

Specification

The title of the invention is not descriptive. Applicant should avoid the use of "novel" in the title, as patents are presumed to be novel and unobvious.

Claim Rejections - 35 USC §§ 101, 112, first paragraph

35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claims 1, 3-18 are rejected under 35 U.S.C. § 101 because they are drawn to an invention with no apparent or disclosed patentable utility. The instant application has provided a description of an isolated DNA encoding a protein and the protein encoded thereby. The instant application does not disclose the biological role of this protein or its significance. The claimed invention is not supported by either a specific and substantial asserted utility or a well

established utility. Novel biological molecules lack well-established utility and must undergo extensive experimentation. Applicant is directed to the Utility Examination Guidelines, Federal Register, Vol. 66, No. 4, pages 1092-1099, Friday January 5, 2001.

It is clear from the instant specification that the nucleic acid encoding the NGPCR polypeptide has been assigned a function because of its similarity to known proteins (Specification at 2, lines 17-23). However, it is commonly known in the art that sequence-to-function methods of assigning protein function are prone to errors (Doerks et al.1998). These errors can be due to sequence similarity of the query region to a region of the alleged similar protein that is not the active site, as well as homologs that did not have the same catalytic activity because active site residues of the characterized family were not conserved (Doerks et al. page 248, column 3, fourth and fifth paragraphs). Inaccurate use of sequence-to-function methods have led to significant function-annotation errors in the sequence databases (Doerks et al. page 250, column 1, third paragraph). Furthermore, Brenner (1999, Trends in Genetics 15:132-133) argues that accurate inference of function from homology must be a difficult problem since, assuming there are only about 1000 major gene superfamilies in nature, then most homologs must have different molecular and cellular functions. Additionally, Bork et al. (1996, Trends in Genetics 12:425-427) add that the software robots that assign functions to new proteins often assign a function to a whole new protein based on structural similarity of a small domain of the new protein to a small domain of a known protein. Such questionable interpretations are written into the sequence database and are then considered facts. Additionally, Yan et al. teaches that in certain cases, a difference of only two-amino acid residues in a protein results in switching the binding of the protein from one receptor to another (Yan et al., Two-amino acid molecular

switch in an epithelial morphogen that regulates binding to two distinct receptors. *Science* 290: 523-527, 2000).

Additionally, even if, *arguendo*, the nucleic acid encoding the NGPCR protein is found to be a G-protein coupled receptor, it is an orphan receptor. Since the ligand to this receptor is unknown, the function of the protein is also unknown. Neither the specification nor the art of record disclose any diseases or conditions associated with the function or expression of the NGPCR protein, therefore, there is no "real world" context of use. Further research to identify or reasonably confirm a "real world" context of use is required. In the instant case, the fact that the claimed invention encodes a GPCR is not sufficient to establish a specific and substantial utility. Although GPCRs have been found to be involved in many different processes and have been the target of much research and drug discovery, unless the specific ligand for each receptor is known, unless the biological activity of the receptor is disclosed and unless the processes that each receptor is involved in are identified, the receptor has no "real world" use, and therefore, lacks specific and substantial utility.

The specification that the nucleic acid of the instant application can be used in screening assays to identify agents which modulate NGPCR receptor signal activity, NGPCR ligands, or levels of mRNA encoding NGPCR (Specification at 45). However, this asserted utility is not specific or substantial. Such assays can be performed with any polynucleotide. Nothing is disclosed about how the polynucleotide is affected by the compounds, which in turn affect production of mRNA and polypeptide. Additionally, the specification discloses nothing specific or substantial for the mRNA and polypeptide produced in this method. Since this asserted utility

is also not present in mature form, so that it could be readily used in a real world sense, the asserted utility is not substantial.

The Specification asserts that the polynucleotide of the instant application can be used in a gene chip to measure expression (Specification at 11). However, this asserted utility is credible but not specific or substantial. Such assays can be performed with any polynucleotide. Further, the specification does not disclose the tissues or cell types the polypeptide/mRNA are normally expressed in. The specification also discloses nothing about the normal levels of expression of the polypeptide/mRNA. The abnormal levels of the polypeptide/mRNA cannot be determined until a baseline control level is established. Applicant further argues that the instant polynucleotides can be used in genome mapping. This asserted utility is credible but not specific or substantial. Such assays can be performed with any polynucleotide. Further, the specification does not disclose a specific DNA target.

After complete characterization, the polynucleotide may be found to encode a polypeptide that has a patentable utility. This further characterization, however, is part of the act of invention and until it has been undertaken Applicant's claimed invention is incomplete. The instant situation is directly analogous to that which was addressed in *Brenner v. Manson,* 148 USPQ 689 (Sup. Ct., 1966), in which a novel compound which was structurally analogous to other compounds which were known to possess anticancer activity was alleged to be potentially useful as an antitumor agent in the absence of evidence supporting this utility. The court expressed the opinion that all chemical compounds are "useful" to the chemical arts when this term is given its broadest interpretation. However, the court held that this broad interpretation was not the intended definition of "useful" as it appears in 35 USC § 101, which requires that an

invention must have either an immediately obvious or fully disclosed "real world" utility. The court held that:

> "The basic quid pro quo contemplated by the Constitution and the Congress for granting a patent monopoly is the benefit derived by the public from an invention with substantial utility", "[u]nless and until a process is refined and developed to this point-where specific benefit exists in currently available form-there is insufficient justification for permitting an applicant to engross what may prove to be a broad field", and "a patent is not a hunting license", "[i]t is not a reward for the search, but compensation for its successful conclusion."

The instant claims are drawn to a nucleic acid encoding a polypeptide which has an as yet undetermined function or biological significance. Until some actual and specific significance can be attributed to the protein identified in the specification as NGPCR, the instant invention is incomplete. The polypeptide encoded by the nucleic acids of the instant invention is known to be structurally analogous to proteins that are known in the art as G protein coupled receptors. In the absence of knowledge of the natural substrate or biological significance of this protein, there is no immediately obvious patentable use for it. To employ a protein of the instant invention in the identification of substances which inhibit its activity is clearly to use it as the object of further research which has been determined by the courts to be a non-patentable utility. Since the instant specification does not disclose a "real world" use for NGPCR then the claimed invention is incomplete and, therefore, does not meet the requirements of 35 USC § 101 as being useful.

Claims 1, 3-18 are also rejected under 35 U.S.C. 112, first paragraph. Specifically, since the claimed invention is not supported by either a specific and substantial asserted utility or a well established utility for the reasons set forth above, one skilled in the art clearly would not know how to use the claimed invention.

Even if, *arguendo*, a patentable utility is found for SEQ ID NO: 1 and 3, claims 1, 3 are rejected under 35 U.S.C. 112, first paragraph, because the specification, which would be enabling for a for a full-length NGPCR polynucleotide of SEQ ID NO: 1 or 3, does not reasonably provide enablement for a nucleic acid sequence comprising at least 24 contiguous nucleotides of SEQ ID NO: 1. The specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the invention commensurate in scope with these claims.

Claims 1, 3 are overly broad since insufficient guidance is provided as to which of the myriad of variant polynucleotides will encode polypeptides will retain the characteristics of NGPCR. Applicants do not disclose any actual or prophetic examples on expected performance parameters of any of the possible variants of NGPCR. It is known in the art that even single amino acid changes or differences in the amino acid sequence of a protein can have dramatic effects on the protein's function. As an example of the unpredictable effects of mutations on protein function, Mickle et al. (Mickle JE et al. Genotype-phenotype relationships in cystic fibrosis. Med Clin North Am. 2000 May;84(3):597-607) teaches that cystic fibrosis is an autosomal recessive disorder caused by abnormal function of a chloride channel, referred to as the cystic fibrosis transmembrane conductance regulator (CFTR) (page 597). Several mutations can cause CF, including the G551D mutation. In this mutation a glycine replaces the aspartic

acid at position 551, giving rise to the CF phenotype. In the most common CF mutation, delta-F508, a single phenylalanine is deleted at position 508, giving ride to the CF phenotype. Thus showing that even the substitution or deletion of a single amino acid in the entire 1480 amino acid CFTR protein sequence can have dramatic and unpredictable effects on the function of the protein. Additionally, it is known in the art that even a single amino acid change in a protein's sequence can drastically affect the structure of the protein and the architecture of an entire cell. For example, Voet et al. (Voet et al. Biochemistry. 1990. John Wiley & Sons, Inc. pages 126-128 and 228-234) teaches that a single Glu to Val substitution in the beta subunit of hemoglobin causes the hemoglobin molecules to associate with one another in such a manner that, in homozygous individuals, erythrocytes are altered from their normal discoid shape and assume the sickle shape characteristic of sickle-cell anemia, causing hemolytic anemia and blood flow blockages (pages 126-128, section 6-3A and page 230, column 2, first paragraph). Additionally, as set foth above, Yan et al. teaches that in certain cases, a change of only two-amino acid residues in a protein results in switching the binding of the protein from one receptor to another (Yan et al., Two-amino acid molecular switch in an epithelial morphogen that regulates binding to two distinct receptors. *Science* 290: 523-527, 2000). Since the claims encompass variant polypeptides and given the art recognized unpredictability of the effect of mutations on protein function, it would require undue experimentation to make and use the claimed invention. See In re Wands, 858 F.2d at 737, 8 USPQ2d at 1404. The test of enablement is not whether any experimentation is necessary, but whether, if experimentation is necessary, it is undue. The claims do not set forth a functional limitation for the encoded polypeptide. Additionally, the amino acid sequence of a polypeptide determines its structural and functional properties, and the

predictability of which amino acids can be substituted is extremely complex and outside the realm of routine experimentation, because accurate predictions of a polypeptide's structure from mere sequence data are limited. Since detailed information regarding the structural and functional requirements of the polynucleotide and the encoded polypeptide are lacking, it is unpredictable as to which variations, if any, meet the limitations of the claims. Applicant is required to enable one of skill in the art to make and use the claimed invention, while the claims encompass polynucleotides and encoded polypeptides which the specification only teaches one skilled in the art to test for functional variants. It would require undue experimentation for one of skill in the art to make and use the claimed polypeptides. Since the claims do not enable one of skill in the art to make and use the claimed polypeptides, but only teaches how to screen for the claimed polypeptides, and since detailed information regarding the structural and functional requirements of the polypeptides are lacking, it is unpredictable as to which variations, if any, meet the limitations of the claims. Thus, since Applicant has only taught how to test for polynucleotides encoding polypeptide variants of NGPCR, and has not taught how to make polynucleotides encoding polypeptide variants of NGPCR, it would require undue experimentation of one of skill in the art to make and use the claimed polynucleotides.

Claims 1, 3 are rejected, under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. Applicant is directed to the Guidelines for the Examination

of Patent Applications Under the 35 U.S.C. 112, ¶ 1 "Written Description" Requirement, Federal Register, Vol. 66, No. 4, pages 1099-1111, Friday January 5, 2001.

The claims are drawn to a nucleic acid sequence comprising at least 24 contiguous nucleotides of SEQ ID NO: 1. These are genus claims because the claims are thus directed to polynucleotides encoding variant polypeptides. The specification and claim do not indicate what distinguishing attributes shared by the members of the genus. The scope of the claim includes numerous structural variants, and the genus is highly variant because a significant number of structural differences between genus members is permitted. The specification and claim do not provide any guidance as to what changes should be made. Structural features that could distinguish compounds in the genus from others in the protein class are missing from the disclosure. No common structural attributes identify the members of the genus. The general knowledge and level of skill in the art do not supplement the omitted description because specific, not general, guidance is what is needed. Since the disclosure fails to describe the common attributes or characteristics that identify members of the genus, and because the genus is highly variant, SEQ ID NO: 1and 3 are insufficient to describe the genus. The written description requirement for a claimed genus may be satisfied through sufficient description of a representative number of species by actual reduction to practice, reduction to drawings, or by disclosure of relevant identifying characteristics, i.e. structure or other physical and/or chemical properties, by functional characteristics coupled with a known or disclosed correlation between structure and function structure, or by a combination of such identifying characteristics, sufficient to show the applicant was in possession of the claimed genus. In the instant case, the specification fails to provide sufficient descriptive information, such as definitive structural or

functional features of the genus of polypeptides. There is no description of the conserved regions that are critical to the structure and function of the genus claimed. There is no description of the sites at which variability may be tolerated and there is no information regarding the relation of structure to function. Structural features that could distinguish the compounds in the genus from other seven transmembrane region compounds are missing from the disclosure. Furthermore, the prior art does not provide compensatory structural or correlative teachings sufficient to enable one of skill to isolate and identify the polynucleotides and polypeptides encompassed. Thus, no identifying characteristics or properties of the instant polypeptides are provided such that one of skill would be able to predictably identify the encompassed molecules as being identical to those instantly claimed. One of skill in the art would reasonably conclude that the disclosure fails to provide a representative number of species to describe the genus. Thus, applicant was not in possession of the claimed genus.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

Claims 1, 3 are rejected under 35 U.S.C. 102(a) as being anticipated by Corby (2000).

The claims are drawn to as nucleic acid molecule comprising at least 24 contiguous nucleotides from SEQ ID NO: 1, and further wherein the nucleic acid is cDNA. The Corby

reference teaches a nucleic acid that is 74.0% identical to SEQ ID NO: 1 (see Sequence Comparison A, attached), and this sequence comprises more than 24 contiguous nucleotides of SEQ ID NO: 1, and the nucleic acid was cloned from cDNA, thus claims 1 and 3 are anticipated.

Conclusion

No claim is allowed.

Advisory Information

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Joseph Murphy whose telephone number is (571) 272-0877. The examiner can normally be reached Monday through Friday from 7:30 am to 5:00 pm. A message may be left on the examiner's voice mail service. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Kunz can be reached on (571) 272-0887.

The fax number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Joseph F. Murphy, Ph. D.
Patent Examiner
Art Unit 1646
June 29, 2004

10008574 Results

SEQ ID NO: 1

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1923	100.0	1923	6	AX686770	AX686770 Sequence
2	1923	100.0	2166	6	AX686774	AX686774 Sequence
3	1915.2	99.6	1920	6	BD182002	BD182002 Novel G p
4	1859.2	96.7	1971	6	BD181999	BD181999 Novel G p
5	1846.4	96.0	2127	9	AY140953	AY140953 Homo sapi
6	1812.6	94.3	1912	6	AX451921	AX451921 Sequence
7	1737	90.3	1737	6	AX686772	AX686772 Sequence
8	1666.8	86.7	2088	6	AX411548	AX411548 Sequence
9	1422.8	74.0	4213	6	AX646687	AX646687 Sequence
10	1422.8	74.0	4213	9	AB065684	AB065684 Homo sapi
c 11	1422.8	74.0	152036	2	AL161776	AL161776 Homo sapi
12	1422.8	74.0	170532	9	AL356421	AL356421 Human DNA
13	1247.8	64.9	1251	6	BD144291	BD144291 Novel G-p
14	1247.8	64.9	1251	9	AB083617	AB083617 Homo sapi
15	974.6	50.7	180643	2	AC117257	AC117257 Mus muscu
c 16	956.6	49.7	177174	2	AC120281	AC120281 Rattus no

RESULT 12
AL356421
LOCUS AL356421 170532 bp DNA linear PRI 30-SEP-2000
DEFINITION Human DNA sequence from clone RP11-550C4 on chromosome 6, complete sequence.
ACCESSION AL356421
VERSION AL356421.10 GI:10443437
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 170532)
AUTHORS Corby,N.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Oct 1, 2000 this sequence version replaced gi:10186530.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6
RP11-550C4 is from the library RPCI-11.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see http://bacpac.med.buffalo.edu/

```
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-550C4 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true left end of clone RP11-550C4 is at 1 in this sequence. The
            true left end of clone RP3-402H5 is at 170433 in this sequence. The
            true right end of clone RP11-812I20 is at 111382 in this sequence.
FEATURES             Location/Qualifiers
     source          1. .170532
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-550C4"
                     /clone_lib="RPCI-11.2"
     misc_feature    41769
                     /note="Tandem repeat. Forced join. Gap size estimated to
                     be approximately 150bp by EcoRI and HindIII restriction
                     enzyme digest data."
BASE COUNT    49941 a  32453 c  34166 g  53972 t
ORIGIN

  Query Match             74.0%;  Score 1422.8;  DB 9;  Length 170532;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1424;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        498 GCAGAGTTACAGCACCATAGCCAACCACATTCTTAACAGCAAAAGCATCTCCAACTGGAC 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56950 GCAGAGTTACAGCACCATAGCCAACCACATTCTTAACAGCAAAAGCATCTCCAACTGGAC 57009

Qy        558 TTTCATTCCTGACAGAAACAGCAGCTATATCCTGCTACATTCAGTCAACTCCTTTGCAAG 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57010 TTTCATTCCTGACAGAAACAGCAGCTATATCCTGCTACATTCAGTCAACTCCTTTGCAAG 57069

Qy        618 AAGGCTATTCATAGATAAACATCCTGTTGACATATCAGATGTCTTCATTCATACTATGGG 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57070 AAGGCTATTCATAGATAAACATCCTGTTGACATATCAGATGTCTTCATTCATACTATGGG 57129

Qy        678 CACCACCATATCTGGAGATAACATTGGAAAAAATTTCACTTTTTCTATGAGAATTAATGA 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57130 CACCACCATATCTGGAGATAACATTGGAAAAAATTTCACTTTTTCTATGAGAATTAATGA 57189

Qy        738 TACCAGCAATGAAGTCACTGGGAGAGTGTTGATCAGCAGAGATGAACTTCGGAAGGTGCC 797
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57190 TACCAGCAATGAAGTCACTGGGAGAGTGTTGATCAGCAGAGATGAACTTCGGAAGGTGCC 57249

Qy        798 TTCCCCTTCTCAGGTCATCAGCATTGCATTTCCAACTATTGGGGCTATTTTGGAAGCCAG 857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57250 TTCCCCTTCTCAGGTCATCAGCATTGCATTTCCAACTATTGGGGCTATTTTGGAAGCCAG 57309

Qy        858 TCTTTTGGAAAATGTTACTGTAAATGGGCTTGTCCTGTCTGCCATTTTGCCCAAGGAACT 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57310 TCTTTTGGAAAATGTTACTGTAAATGGGCTTGTCCTGTCTGCCATTTTGCCCAAGGAACT 57369

Qy        918 TAAAAGAATCTCACTGATTTTTGAAAAGATCAGCAAGTCAGAGGAGAGGAGGACACAGTG 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57370 TAAAAGAATCTCACTGATTTTTGAAAAGATCAGCAAGTCAGAGGAGAGGAGGACACAGTG 57429

Qy        978 TGTTGGCTGGCACTCTGTGGAGAACAGATGGGACCAGCAGGCCTGCAAAATGATTCAAGA 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57430 TGTTGGCTGGCACTCTGTGGAGAACAGATGGGACCAGCAGGCCTGCAAAATGATTCAAGA 57489

Qy       1038 AAACTCCCAGCAAGCTGTTTGCAAATGTAGGCCAAGTGAATTGTTTACCTCTTTCTCAAT 1097
              |||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db      57490 AAACTCCCAGCAAGCTGTTTGCAAATGTAGGCCAAGCAAATTGTTTACCTCTTTCTCAAT 57549

Qy       1098 TCTTATGTCACCTCACATCTTAGAGAGTCTGATTCTGACTTACATCACATATGTAGGCCT 1157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57550 TCTTATGTCACCTCACATCTTAGAGAGTCTGATTCTGACTTACATCACATATGTAGGCCT 57609

Qy       1158 GGGCATTTCTATTTGCAGCCTGATCCTTTGCTTGTCCATTGAGGTCCTAGTCTGGAGCCA 1217
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57610 GGGCATTTCTATTTGCAGCCTGATCCTTTGCTTGTCCATTGAGGTCCTAGTCTGGAGCCA 57669

Qy       1218 AGTGACAAAGACAGAGATCACCTATTTACGCCATGTGTGCATTGTTAACATTGCAGCCAC 1277
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57670 AGTGACAAAGACAGAGATCACCTATTTACGCCATGTGTGCATTGTTAACATTGCAGCCAC 57729

Qy       1278 TTTGCTGATGGCAGATGTGTGGTTCATTGTGGCTTCCTTTCTTAGTGGCCCAATAACACA 1337
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57730 TTTGCTGATGGCAGATGTGTGGTTCATTGTGGCTTCCTTTCTTAGTGGCCCAATAACACA 57789

Qy       1338 CCACAAGGGATGTGTGGCAGCCACATTTTTTGTTCATTTCTTTTACCTTTCTGTATTTTT 1397
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57790 CCACAAGGGATGTGTGGCAGCCACATTTTTTGTTCATTTCTTTTACCTTTCTGTATTTTT 57849

Qy       1398 CTGGATGCTTGCCAAGGCACTCCTTATCCTCTATGGAATCATGATTGTTTTCCATACCTT 1457
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57850 CTGGATGCTTGCCAAGGCACTCCTTATCCTCTATGGAATCATGATTGTTTTCCATACCTT 57909

Qy       1458 GCCCAAGTCAGTCCTGGTGGCATCTCTGTTTTCAGTGGGCTATGGATGCCCTTTGGCCAT 1517
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57910 GCCCAAGTCAGTCCTGGTGGCATCTCTGTTTTCAGTGGGCTATGGATGCCCTTTGGCCAT 57969

Qy       1518 TGCTGCCATCACTGTTGCTGCCACTGAACCTGGCAAAGGCTATCTACGACCTGAGATCTG 1577
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57970 TGCTGCCATCACTGTTGCTGCCACTGAACCTGGCAAAGGCTATCTACGACCTGAGATCTG 58029

Qy       1578 CTGGCTCAACTGGGACATGACCAAAGCCCTCCTGGCCTTCGTGATCCCAGCTTTGGCCAT 1637
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58030 CTGGCTCAACTGGGACATGACCAAAGCCCTCCTGGCCTTCGTGATCCCAGCTTTGGCCAT 58089

Qy       1638 CGTGGTAGTAAACCTGATCACAGTCACACTGGTGATTGTCAAGACCCAGCGAGCTGCCAT 1697
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58090 CGTGGTAGTAAACCTGATCACAGTCACACTGGTGATTGTCAAGACCCAGCGAGCTGCCAT 58149

Qy       1698 TGGCAATTCCATGTTCCAGGAAGTGAGAGCCATTGTGAGAATCAGCAAGAACATCGCCAT 1757
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58150 TGGCAATTCCATGTTCCAGGAAGTGAGAGCCATTGTGAGAATCAGCAAGAACATCGCCAT 58209

Qy       1758 CCTCACACCACTTCTGGGACTGACCTGGGGATTTGGAGTAGCCACTGTCATCGATGACAG 1817
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58210 CCTCACACCACTTCTGGGACTGACCTGGGGATTTGGAGTAGCCACTGTCATCGATGACAG 58269

Qy       1818 ATCCCTGGCCTTCCACATTATCTTCTCCCTGCTCAATGCATTCCAGGTAAGTCCAGATGC 1877
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58270 ATCCCTGGCCTTCCACATTATCTTCTCCCTGCTCAATGCATTCCAGGTAAGTCCAGATGC 58329

Qy       1878 TTCTGACCAAGTGCAAAGTGAGAGAATTCATGAAGATGTTCTGTGA 1923
                ||||||||||||||||||||||||||||||||||||||||||||
Db      58330 TTCTGACCAAGTGCAAAGTGAGAGAATTCATGAAGATGTTCTGTGA 58375
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1923	100.0	1923	24	ABA00446	Human GPCR cDNA #1
2	1923	100.0	2166	24	ABA00448	Human GPCR ORF and
3	1915.2	99.6	1920	25	ABZ24092	Human GPCR protein
4	1859.2	96.7	1971	25	ABZ24089	Human GPCR protein
5	1812.6	94.3	1912	24	AAD37666	Human G-protein co
6	1737	90.3	1737	24	ABA00447	Human GPCR cDNA #2
7	1666.8	86.7	2088	24	ABN88263	Human secretin rec
8	1247.8	64.9	1251	24	ABZ42885	Human GPCR polynuc
9	572.2	29.8	1971	24	ABK49800	Human cDNA encodin
10	570.8	29.7	2112	24	ABL60552	Human secretin rec
11	568.4	29.6	2094	24	ABL60558	Human secretin rec
12	567.2	29.5	3230	25	ABZ59302	Human GPCR clone 1
13	566.8	29.5	2085	24	ABK49803	Human cDNA encodin
14	566.8	29.5	3410	25	AAD50425	Human GPCR cDNA.

15	565.2	29.4	2322	24	AAD29679	Human G-protein co
16	543.4	28.3	1527	24	ABK49808	Human cDNA encodin
17	519	27.0	17198	25	AAD50426	Human GPCR gene.
18	509	26.5	1626	22	AAF28687	Human protein HP10
19	509	26.5	2667	22	AAF28697	Human protein HP10
20	509	26.5	2667	25	ABZ42813	Human G protein-co

SEQ ID NO: 1 oligo 24

Result No.	Score	Query Match	Length	DB	ID	Description
1	1923	100.0	1923	6	AX686770	AX686770 Sequence
2	1923	100.0	2166	6	AX686774	AX686774 Sequence
3	1767	91.9	1920	6	BD182002	BD182002 Novel G p
4	1737	90.3	1737	6	AX686772	AX686772 Sequence
5	1711	89.0	1971	6	BD181999	BD181999 Novel G p
6	1697	88.2	2127	9	AY140953	AY140953 Homo sapi
7	1371	71.3	2088	6	AX411548	AX411548 Sequence
8	1370	71.2	1912	6	AX451921	AX451921 Sequence
9	1324	68.9	4213	6	AX646687	AX646687 Sequence
10	1324	68.9	4213	9	AB065684	AB065684 Homo sapi
c 11	1324	68.9	152036	2	AL161776	AL161776 Homo sapi
12	1324	68.9	170532	9	AL356421	AL356421 Human DNA
13	1149	59.8	1251	6	BD144291	BD144291 Novel G-p
14	1149	59.8	1251	9	AB083617	AB083617 Homo sapi
15	384	20.0	486	9	AY255612	AY255612 Homo sapi
16	228	11.9	330	6	AX147792	AX147792 Sequence
17	228	11.9	330	6	AX521841	AX521841 Sequence
18	27	1.4	180643	2	AC117257	AC117257 Mus muscu
19	26	1.4	26	6	BD182000	BD182000 Novel G p
20	26	1.4	26	6	BD182006	BD182006 Novel G p
c 21	25	1.3	25	6	BD182003	BD182003 Novel G p

RESULT 12
AL356421
LOCUS AL356421 170532 bp DNA linear PRI 30-SEP-2000
DEFINITION Human DNA sequence from clone RP11-550C4 on chromosome 6, complete sequence.
ACCESSION AL356421
VERSION AL356421.10 GI:10443437
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 170532)
AUTHORS Corby,N.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Oct 1, 2000 this sequence version replaced gi:10186530.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.
The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6
RP11-550C4 is from the library RPCI-11.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see http://bacpac.med.buffalo.edu/
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone RP11-550C4 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true left end of clone RP11-550C4 is at 1 in this sequence. The true left end of clone RP3-402H5 is at 170433 in this sequence. The true right end of clone RP11-812I20 is at 111382 in this sequence.

```
FEATURES             Location/Qualifiers
     source          1. .170532
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-550C4"
                     /clone_lib="RPCI-11.2"
     misc_feature    41769
                     /note="Tandem repeat. Forced join. Gap size estimated to
                     be approximately 150bp by EcoRI and HindIII restriction
                     enzyme digest data."
BASE COUNT    49941 a  32453 c  34166 g  53972 t
ORIGIN

  Query Match             68.9%;  Score 1324;  DB 9;  Length 170532;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1424;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        498 GCAGAGTTACAGCACCATAGCCAACCACATTCTTAACAGCAAAAGCATCTCCAACTGGAC 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56950 GCAGAGTTACAGCACCATAGCCAACCACATTCTTAACAGCAAAAGCATCTCCAACTGGAC 57009

Qy        558 TTTCATTCCTGACAGAAACAGCAGCTATATCCTGCTACATTCAGTCAACTCCTTTGCAAG 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57010 TTTCATTCCTGACAGAAACAGCAGCTATATCCTGCTACATTCAGTCAACTCCTTTGCAAG 57069

Qy        618 AAGGCTATTCATAGATAAACATCCTGTTGACATATCAGATGTCTTCATTCATACTATGGG 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57070 AAGGCTATTCATAGATAAACATCCTGTTGACATATCAGATGTCTTCATTCATACTATGGG 57129

Qy        678 CACCACCATATCTGGAGATAACATTGGAAAAAATTTCACTTTTTCTATGAGAATTAATGA 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57130 CACCACCATATCTGGAGATAACATTGGAAAAAATTTCACTTTTTCTATGAGAATTAATGA 57189

Qy        738 TACCAGCAATGAAGTCACTGGGAGAGTGTTGATCAGCAGAGATGAACTTCGGAAGGTGCC 797
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57190 TACCAGCAATGAAGTCACTGGGAGAGTGTTGATCAGCAGAGATGAACTTCGGAAGGTGCC 57249

Qy        798 TTCCCCTTCTCAGGTCATCAGCATTGCATTTCCAACTATTGGGGCTATTTTGGAAGCCAG 857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57250 TTCCCCTTCTCAGGTCATCAGCATTGCATTTCCAACTATTGGGGCTATTTTGGAAGCCAG 57309

Qy        858 TCTTTTGGAAAATGTTACTGTAAATGGGCTTGTCCTGTCTGCCATTTTGCCCAAGGAACT 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57310 TCTTTTGGAAAATGTTACTGTAAATGGGCTTGTCCTGTCTGCCATTTTGCCCAAGGAACT 57369

Qy        918 TAAAAGAATCTCACTGATTTTTGAAAAGATCAGCAAGTCAGAGGAGAGGAGGACACAGTG 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57370 TAAAAGAATCTCACTGATTTTTGAAAAGATCAGCAAGTCAGAGGAGAGGAGGACACAGTG 57429

Qy        978 TGTTGGCTGGCACTCTGTGGAGAACAGATGGGACCAGCAGGCCTGCAAAATGATTCAAGA 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      57430 TGTTGGCTGGCACTCTGTGGAGAACAGATGGGACCAGCAGGCCTGCAAAATGATTCAAGA 57489

Qy       1038 AAACTCCCAGCAAGCTGTTTGCAAATGTAGGCCAAGTGAATTGTTTACCTCTTTCTCAAT 1097
              ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db      57490 AAACTCCCAGCAAGCTGTTTGCAAATGTAGGCCAAGCAAATTGTTTACCTCTTTCTCAAT 57549

Qy       1098 TCTTATGTCACCTCACATCTTAGAGAGTCTGATTCTGACTTACATCACATATGTAGGCCT 1157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57550 TCTTATGTCACCTCACATCTTAGAGAGTCTGATTCTGACTTACATCACATATGTAGGCCT 57609

Qy       1158 GGGCATTTCTATTTGCAGCCTGATCCTTTGCTTGTCCATTGAGGTCCTAGTCTGGAGCCA 1217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57610 GGGCATTTCTATTTGCAGCCTGATCCTTTGCTTGTCCATTGAGGTCCTAGTCTGGAGCCA 57669

Qy       1218 AGTGACAAAGACAGAGATCACCTATTTACGCCATGTGTGCATTGTTAACATTGCAGCCAC 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57670 AGTGACAAAGACAGAGATCACCTATTTACGCCATGTGTGCATTGTTAACATTGCAGCCAC 57729

Qy       1278 TTTGCTGATGGCAGATGTGTGGTTCATTGTGGCTTCCTTTCTTAGTGGCCCAATAACACA 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57730 TTTGCTGATGGCAGATGTGTGGTTCATTGTGGCTTCCTTTCTTAGTGGCCCAATAACACA 57789

Qy       1338 CCACAAGGGATGTGTGGCAGCCACATTTTTTGTTCATTTCTTTTACCTTTCTGTATTTTT 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57790 CCACAAGGGATGTGTGGCAGCCACATTTTTTGTTCATTTCTTTTACCTTTCTGTATTTTT 57849

Qy       1398 CTGGATGCTTGCCAAGGCACTCCTTATCCTCTATGGAATCATGATTGTTTTCCATACCTT 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57850 CTGGATGCTTGCCAAGGCACTCCTTATCCTCTATGGAATCATGATTGTTTTCCATACCTT 57909

Qy       1458 GCCCAAGTCAGTCCTGGTGGCATCTCTGTTTTCAGTGGGCTATGGATGCCCTTTGGCCAT 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57910 GCCCAAGTCAGTCCTGGTGGCATCTCTGTTTTCAGTGGGCTATGGATGCCCTTTGGCCAT 57969

Qy       1518 TGCTGCCATCACTGTTGCTGCCACTGAACCTGGCAAAGGCTATCTACGACCTGAGATCTG 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57970 TGCTGCCATCACTGTTGCTGCCACTGAACCTGGCAAAGGCTATCTACGACCTGAGATCTG 58029

Qy       1578 CTGGCTCAACTGGGACATGACCAAAGCCCTCCTGGCCTTCGTGATCCCAGCTTTGGCCAT 1637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58030 CTGGCTCAACTGGGACATGACCAAAGCCCTCCTGGCCTTCGTGATCCCAGCTTTGGCCAT 58089

Qy       1638 CGTGGTAGTAAACCTGATCACAGTCACACTGGTGATTGTCAAGACCCAGCGAGCTGCCAT 1697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58090 CGTGGTAGTAAACCTGATCACAGTCACACTGGTGATTGTCAAGACCCAGCGAGCTGCCAT 58149

Qy       1698 TGGCAATTCCATGTTCCAGGAAGTGAGAGCCATTGTGAGAATCAGCAAGAACATCGCCAT 1757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58150 TGGCAATTCCATGTTCCAGGAAGTGAGAGCCATTGTGAGAATCAGCAAGAACATCGCCAT 58209

Qy       1758 CCTCACACCACTTCTGGGACTGACCTGGGGATTTGGAGTAGCCACTGTCATCGATGACAG 1817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58210 CCTCACACCACTTCTGGGACTGACCTGGGGATTTGGAGTAGCCACTGTCATCGATGACAG 58269

Qy       1818 ATCCCTGGCCTTCCACATTATCTTCTCCCTGCTCAATGCATTCCAGGTAAGTCCAGATGC 1877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58270 ATCCCTGGCCTTCCACATTATCTTCTCCCTGCTCAATGCATTCCAGGTAAGTCCAGATGC 58329

Qy       1878 TTCTGACCAAGTGCAAAGTGAGAGAATTCATGAAGATGTTCTGTGA 1923
              |||||||||||||||||||||||||||||||||||||||||||||
Db      58330 TTCTGACCAAGTGCAAAGTGAGAGAATTCATGAAGATGTTCTGTGA 58375
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1923	100.0	1923	24	ABA00446	Human GPCR cDNA #1
2	1923	100.0	2166	24	ABA00448	Human GPCR ORF and
3	1767	91.9	1920	25	ABZ24092	Human GPCR protein
4	1737	90.3	1737	24	ABA00447	Human GPCR cDNA #2

5	1711	89.0	1971	25	ABZ24089	Human GPCR protein
6	1371	71.3	2088	24	ABN88263	Human secretin rec
7	1370	71.2	1912	24	AAD37666	Human G-protein co
8	1149	59.8	1251	24	ABZ42885	Human GPCR polynuc
9	228	11.9	330	22	AAH50987	Human nGPCR31 codi
10	228	11.9	330	24	ABS70220	DNA encoding human
11	26	1.4	26	24	ABN88265	Human secretin rec
12	26	1.4	26	25	ABZ24090	Human GPCR protein
13	26	1.4	26	25	ABZ24096	Human GPCR protein
14	25	1.3	25	25	ABZ59314	Human GPCR related
c 15	25	1.3	25	25	ABZ59315	Human GPCR related
c 16	25	1.3	25	25	ABZ24093	Human GPCR protein

SEQ ID NO: 3

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1737	100.0	1737	6	AX686772	AX686772 Sequence
2	1737	100.0	1923	6	AX686770	AX686770 Sequence
3	1737	100.0	2166	6	AX686774	AX686774 Sequence
4	1731.8	99.7	2127	9	AY140953	AY140953 Homo sapi
5	1730.8	99.6	1920	6	BD182002	BD182002 Novel G p
6	1724.2	99.3	1912	6	AX451921	AX451921 Sequence
7	1674.8	96.4	1971	6	BD181999	BD181999 Novel G p
8	1596.4	91.9	2088	6	AX411548	AX411548 Sequence
9	1422.8	81.9	4213	6	AX646687	AX646687 Sequence
10	1422.8	81.9	4213	9	AB065684	AB065684 Homo sapi
c 11	1422.8	81.9	152036	2	AL161776	AL161776 Homo sapi
12	1422.8	81.9	170532	9	AL356421	AL356421 Human DNA
13	1247.8	71.8	1251	6	BD144291	BD144291 Novel G-p
14	1247.8	71.8	1251	9	AB083617	AB083617 Homo sapi
15	974.6	56.1	180643	2	AC117257	AC117257 Mus muscu
c 16	956.6	55.1	177174	2	AC120281	AC120281 Rattus no

RESULT 12
AL356421
LOCUS AL356421 170532 bp DNA linear PRI 30-SEP-2000
DEFINITION Human DNA sequence from clone RP11-550C4 on chromosome 6, complete sequence.
ACCESSION AL356421
VERSION AL356421.10 GI:10443437
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 170532)
AUTHORS Corby,N.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Oct 1, 2000 this sequence version replaced gi:10186530.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure'

feature key.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6
RP11-550C4 is from the library RPCI-11.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see http://bacpac.med.buffalo.edu/
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone RP11-550C4 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true left end of clone RP11-550C4 is at 1 in this sequence. The true left end of clone RP3-402H5 is at 170433 in this sequence. The true right end of clone RP11-812I20 is at 111382 in this sequence.

```
FEATURES             Location/Qualifiers
     source          1. .170532
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-550C4"
                     /clone_lib="RPCI-11.2"
     misc_feature    41769
                     /note="Tandem repeat. Forced join. Gap size estimated to
                     be approximately 150bp by EcoRI and HindIII restriction
                     enzyme digest data."
BASE COUNT    49941 a  32453 c  34166 g  53972 t
ORIGIN
```

Query Match 81.9%; Score 1422.8; DB 9; Length 170532;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1424; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        312 GCAGAGTTACAGCACCATAGCCAACCACATTCTTAACAGCAAAAGCATCTCCAACTGGAC 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56950 GCAGAGTTACAGCACCATAGCCAACCACATTCTTAACAGCAAAAGCATCTCCAACTGGAC 57009

Qy        372 TTTCATTCCTGACAGAAACAGCAGCTATATCCTGCTACATTCAGTCAACTCCTTTGCAAG 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57010 TTTCATTCCTGACAGAAACAGCAGCTATATCCTGCTACATTCAGTCAACTCCTTTGCAAG 57069

Qy        432 AAGGCTATTCATAGATAAACATCCTGTTGACATATCAGATGTCTTCATTCATACTATGGG 491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57070 AAGGCTATTCATAGATAAACATCCTGTTGACATATCAGATGTCTTCATTCATACTATGGG 57129

Qy        492 CACCACCATATCTGGAGATAACATTGGAAAAAATTTCACTTTTTCTATGAGAATTAATGA 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57130 CACCACCATATCTGGAGATAACATTGGAAAAAATTTCACTTTTTCTATGAGAATTAATGA 57189

Qy        552 TACCAGCAATGAAGTCACTGGGAGAGTGTTGATCAGCAGAGATGAACTTCGGAAGGTGCC 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57190 TACCAGCAATGAAGTCACTGGGAGAGTGTTGATCAGCAGAGATGAACTTCGGAAGGTGCC 57249

Qy        612 TTCCCCTTCTCAGGTCATCAGCATTGCATTTCCAACTATTGGGGCTATTTTGGAAGCCAG 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57250 TTCCCCTTCTCAGGTCATCAGCATTGCATTTCCAACTATTGGGGCTATTTTGGAAGCCAG 57309

Qy        672 TCTTTTGGAAAATGTTACTGTAAATGGGCTTGTCCTGTCTGCCATTTTGCCCAAGGAACT 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57310 TCTTTTGGAAAATGTTACTGTAAATGGGCTTGTCCTGTCTGCCATTTTGCCCAAGGAACT 57369

Qy        732 TAAAAGAATCTCACTGATTTTTGAAAAGATCAGCAAGTCAGAGGAGAGGAGGACACAGTG 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57370 TAAAAGAATCTCACTGATTTTTGAAAAGATCAGCAAGTCAGAGGAGAGGAGGACACAGTG 57429
```

```
Qy        792 TGTTGGCTGGCACTCTGTGGAGAACAGATGGGACCAGCAGGCCTGCAAAATGATTCAAGA 851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57430 TGTTGGCTGGCACTCTGTGGAGAACAGATGGGACCAGCAGGCCTGCAAAATGATTCAAGA 57489

Qy        852 AAACTCCCAGCAAGCTGTTTGCAAATGTAGGCCAAGTGAATTGTTTACCTCTTTCTCAAT 911
              |||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db      57490 AAACTCCCAGCAAGCTGTTTGCAAATGTAGGCCAAGCAAATTGTTTACCTCTTTCTCAAT 57549

Qy        912 TCTTATGTCACCTCACATCTTAGAGAGTCTGATTCTGACTTACATCACATATGTAGGCCT 971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57550 TCTTATGTCACCTCACATCTTAGAGAGTCTGATTCTGACTTACATCACATATGTAGGCCT 57609

Qy        972 GGGCATTTCTATTTGCAGCCTGATCCTTTGCTTGTCCATTGAGGTCCTAGTCTGGAGCCA 1031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57610 GGGCATTTCTATTTGCAGCCTGATCCTTTGCTTGTCCATTGAGGTCCTAGTCTGGAGCCA 57669

Qy       1032 AGTGACAAAGACAGAGATCACCTATTTACGCCATGTGTGCATTGTTAACATTGCAGCCAC 1091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57670 AGTGACAAAGACAGAGATCACCTATTTACGCCATGTGTGCATTGTTAACATTGCAGCCAC 57729

Qy       1092 TTTGCTGATGGCAGATGTGTGGTTCATTGTGGCTTCCTTTCTTAGTGGCCCAATAACACA 1151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57730 TTTGCTGATGGCAGATGTGTGGTTCATTGTGGCTTCCTTTCTTAGTGGCCCAATAACACA 57789

Qy       1152 CCACAAGGGATGTGTGGCAGCCACATTTTTTGTTCATTTCTTTTACCTTTCTGTATTTTT 1211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57790 CCACAAGGGATGTGTGGCAGCCACATTTTTTGTTCATTTCTTTTACCTTTCTGTATTTTT 57849

Qy       1212 CTGGATGCTTGCCAAGGCACTCCTTATCCTCTATGGAATCATGATTGTTTTCCATACCTT 1271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57850 CTGGATGCTTGCCAAGGCACTCCTTATCCTCTATGGAATCATGATTGTTTTCCATACCTT 57909

Qy       1272 GCCCAAGTCAGTCCTGGTGGCATCTCTGTTTTCAGTGGGCTATGGATGCCCTTTGGCCAT 1331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57910 GCCCAAGTCAGTCCTGGTGGCATCTCTGTTTTCAGTGGGCTATGGATGCCCTTTGGCCAT 57969

Qy       1332 TGCTGCCATCACTGTTGCTGCCACTGAACCTGGCAAAGGCTATCTACGACCTGAGATCTG 1391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57970 TGCTGCCATCACTGTTGCTGCCACTGAACCTGGCAAAGGCTATCTACGACCTGAGATCTG 58029

Qy       1392 CTGGCTCAACTGGGACATGACCAAAGCCCTCCTGGCCTTCGTGATCCCAGCTTTGGCCAT 1451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58030 CTGGCTCAACTGGGACATGACCAAAGCCCTCCTGGCCTTCGTGATCCCAGCTTTGGCCAT 58089

Qy       1452 CGTGGTAGTAAACCTGATCACAGTCACACTGGTGATTGTCAAGACCCAGCGAGCTGCCAT 1511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58090 CGTGGTAGTAAACCTGATCACAGTCACACTGGTGATTGTCAAGACCCAGCGAGCTGCCAT 58149

Qy       1512 TGGCAATTCCATGTTCCAGGAAGTGAGAGCCATTGTGAGAATCAGCAAGAACATCGCCAT 1571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58150 TGGCAATTCCATGTTCCAGGAAGTGAGAGCCATTGTGAGAATCAGCAAGAACATCGCCAT 58209

Qy       1572 CCTCACACCACTTCTGGGACTGACCTGGGGATTTGGAGTAGCCACTGTCATCGATGACAG 1631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58210 CCTCACACCACTTCTGGGACTGACCTGGGGATTTGGAGTAGCCACTGTCATCGATGACAG 58269

Qy       1632 ATCCCTGGCCTTCCACATTATCTTCTCCCTGCTCAATGCATTCCAGGTAAGTCCAGATGC 1691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58270 ATCCCTGGCCTTCCACATTATCTTCTCCCTGCTCAATGCATTCCAGGTAAGTCCAGATGC 58329

Qy       1692 TTCTGACCAAGTGCAAAGTGAGAGAATTCATGAAGATGTTCTGTGA 1737
              |||||||||||||||||||||||||||||||||||||||||||||
Db      58330 TTCTGACCAAGTGCAAAGTGAGAGAATTCATGAAGATGTTCTGTGA 58375
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1737	100.0	1737	24	ABA00447	Human GPCR cDNA #2

2	1737	100.0	1923	24	ABA00446	Human GPCR cDNA #1
3	1737	100.0	2166	24	ABA00448	Human GPCR ORF and
4	1730.8	99.6	1920	25	ABZ24092	Human GPCR protein
5	1724.2	99.3	1912	24	AAD37666	Human G-protein co
6	1674.8	96.4	1971	25	ABZ24089	Human GPCR protein
7	1596.4	91.9	2088	24	ABN88263	Human secretin rec
8	1247.8	71.8	1251	24	ABZ42885	Human GPCR polynuc
9	549.8	31.7	1971	24	ABK49800	Human cDNA encodin
10	548.4	31.6	2112	24	ABL60552	Human secretin rec
11	546	31.4	2094	24	ABL60558	Human secretin rec
12	544.8	31.4	3230	25	ABZ59302	Human GPCR clone 1
13	544.4	31.3	2085	24	ABK49803	Human cDNA encodin
14	544.4	31.3	3410	25	AAD50425	Human GPCR cDNA.
15	543.4	31.3	1527	24	ABK49808	Human cDNA encodin
16	542.8	31.2	2322	24	AAD29679	Human G-protein co